# Remarks

In the present response, claims 1-20 are presented for examination.

## Claim Rejections: 35 USC § 103(a)

Claims 1-20 are rejected under 35 USC § 103(a) as being unpatentable over US publication number 2002/0194000 (Bennett) in view of USPN 7,058,573 (Murveit). These rejections are traversed.

Each of the independent claims recites one or more elements that are not taught or suggested in Bennett in view of Murveit. These missing elements show that the differences between the combined teachings in the art and the recitations in the claims are great. As such, the pending claims are <u>not</u> a predictable variation of the art to one of ordinary skill in the art.

Some examples of these missing elements are provided below with respect to the independent claims.

Claim 1

As one example, independent claim 1 recites analyzing characteristics of the first speech utterance to determine an age and a gender of the first user. The utterance is then classified into a category based on the age and the gender of the user. Bennett in view of Murveit does not teach these elements.

Bennett analyzes the input stream (i.e., speech utterance) for context characteristics such as background noise, signal strength, or callerID (see paragraph [0015] in Bennett). Bennett repeatedly expresses the importance of using <u>contextual</u> information about the incoming stream as opposed to content information of the stream: "A more complex characteristic of the incoming stream is contextual information. Contextual information is that information related to the environment around the input stream…" (see paragraph [0018] in Bennett).

Bennett expressly teaches using contextual information to classify the incoming stream. Bennett does not teach or even suggest analyzing characteristics of the first speech utterance to determine an age and a gender of the first user. Murveit also does not

teach or suggest analyzing characteristics of the first speech utterance to determine an age and a gender of the first user.

For at least these reasons, independent claim 1 and its dependent claims are allowable over Bennett in view of Murveit.

As another example, claim 1 recites assigning a different ASR engine to each category based on the ranks of the ASR engines to develop a ranking matrix. The claim then recites consulting the ranking matrix to select an ASR engine. Bennett in view of Murveit does not teach these elements.

In Bennett, the input speech utterance is automatically sent to different recognizers (see paragraph [0010] in Bennett). The system in Bennett then uses a predictor to select the best results from the plural recognizers. The architecture in Bennett (see Figure 1) is setup for using multiple outputs from multiple recognizers to the predictor. Bennett never suggests that his system could somehow use a ranking matrix for selecting one of the recognizers. Bennett appears to teach away from such a matrix since the input speech utterance is advantageously sent to multiple recognizers in Bennett.

For at least these reasons, independent claim 1 and its dependent claims are allowable over Bennett in view of Murveit.


Claim 8

As one example, independent claim 8 recites means for extracting an age and a gender of the user from characteristics of the utterance of the user. Bennett in view of Murveit does not teach this element.

Bennett analyzes the input stream (i.e., speech utterance) for context characteristics such as background noise, signal strength, or callerID (see paragraph [0015] in Bennett). Bennett repeatedly expresses the importance of using contextual information about the incoming stream as opposed to content information of the stream: "A more complex characteristic of the incoming stream is contextual information. Contextual information is that information related to the environment around the input stream…" (see paragraph [0018] in Bennett).

Bennett expressly teaches using contextual information to classify the incoming stream. Bennett does not teach or even suggest extracting an age and a gender of the user

from characteristics of the utterance of the user. Murveit also does not teach or suggest extracting an age and a gender of the user from characteristics of the utterance of the user.

For at least these reasons, independent claim 8 and its dependent claims are allowable over Bennett in view of Murveit.

As another example, claim 8 recites means for comparing output from each of the different ASR engines with the ground truths to develop a ranking matrix that includes ranks of the different ASR engines for accuracy in recognizing the different speech utterances. The claim then recites consulting the ranking matrix to select a best performing ASR engine.

In Bennett, the input speech utterance is automatically sent to different recognizers (see paragraph [0010] in Bennett). The system in Bennett then uses a predictor to select the best results from the plural recognizers. The architecture in Bennett (see Figure 1) is setup for using multiple outputs from multiple recognizers to the predictor. Bennett never suggests that his system could somehow use a ranking matrix for selecting one of the recognizers. Bennett appears to teach away from such a matrix since the input speech utterance is advantageously sent to multiple recognizers in Bennett.

For at least these reasons, independent claim 8 and its dependent claims are allowable over Bennett in view of Murveit.

As another example, claim 8 recites means for using the characteristics to select a best performing ASR engine from the different ASR engines. By contrast, Bennett selects multiple recognizers and then determines which one of the recognizers provides the best estimation of accuracy (see paragraph [0022] in Bennett). The process described in Bennett requires a larger amount of processing power because each speech utterance is sent to multiple recognizers. By contrast, claim 8 recites that a single best ASR engine is selected from plural ASR engines.

For at least these reasons, independent claim 8 and its dependent claims are allowable over Bennett in view of Murveit.

Claim 14

As one example, independent claim 14 recites a computer system that analyzes characteristics of the speech utterance to determine an age and a gender of the speaker. Bennett in view of Murveit does not teach this element.

Bennett analyzes the input stream (i.e., speech utterance) for context characteristics such as background noise, signal strength, or callerID (see paragraph [0015] in Bennett). Bennett repeatedly expresses the importance of using contextual information about the incoming stream as opposed to content information of the stream: "A more complex characteristic of the incoming stream is contextual information. Contextual information is that information related to the environment around the input stream…" (see paragraph [0018] in Bennett).

Bennett expressly teaches using contextual information to classify the incoming stream. Bennett does not teach or even suggest analyzing characteristics of the speech utterance to determine an age and a gender of the speaker. Murveit also does not teach or suggest analyzing characteristics of the speech utterance to determine an age and a gender of the speaker.

For at least these reasons, independent claim 14 and its dependent claims are allowable over Bennett in view of Murveit.

As another example, claim 14 recites comparing output from each of the different ASR engines with the ground truths to develop a ranking matrix that includes ranks of the different ASR engines for accuracy in recognizing the different speech utterances. The claim then recites using the age and gender of the user to select from the ranking matrix one of the ASR engines. Bennett in view of Murveit does not teach these elements.

In Bennett, the input speech utterance is automatically sent to different recognizers (see paragraph [0010] in Bennett). The system in Bennett then uses a predictor to select the best results from the plural recognizers. The architecture in Bennett (see Figure 1) is setup for using multiple outputs from multiple recognizers to the predictor. Bennett never suggests that his system could somehow use a ranking matrix for selecting one of the recognizers. Bennett appears to teach away from such a matrix since the input speech utterance is advantageously sent to multiple recognizers in Bennett.

For at least these reasons, independent claim 14 and its dependent claims are allowable over Bennett in view of Murveit.

As another example, claim 14 recites using the extracted characteristics to select "one of the ASR engines" that most accurately recognizes the speech utterance. By contrast, Bennett selects multiple recognizers and then determines which one of the recognizers provides the best estimation of accuracy (see paragraph [0022] in Bennett). The process described in Bennett requires a larger amount of processing power because each speech utterance is sent to multiple recognizers. By contrast, claim 14 recites that a single ASR engine that will most accurately recognize the speech is selected from plural ASR engines.

For at least these reasons, independent claim 14 and its dependent claims are allowable over Bennett in view of Murveit.

# CONCLUSION

In view of the above, Applicants believe that all pending claims are in condition for allowance. Allowance of these claims is respectfully requested.

Any inquiry regarding this Amendment and Response should be directed to Philip S. Lyren at Telephone No. 832-236-5529. In addition, all correspondence should continue to be directed to the following address:

**Hewlett-Packard Company**
Intellectual Property Administration
P.O. Box 272400
Fort Collins, Colorado 80527-2400

Respectfully submitted,

/Philip S. Lyren #40,709/

Philip S. Lyren
Reg. No. 40,709
Ph: 832-236-5529